JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 56, 502–528 (1976)

Constrained Control Problems with Convex Cost in Hilbert Space

VIOREL BARBU

Faculty of Mathematics, University of Iaşi, Iaşi, Romania

Submitted by J. L. Lions

INTRODUCTION

In recent papers [3, 4] necessary and sufficient conditions were derived for convex control problems involving linear differential equations in Hilbert spaces. In this paper we show that, under relatively weak assumptions (see Section 1), one can derive the above mentioned optimality conditions, for a wide class of control problems for linear evolution equations in Hilbert space.

The plan of the paper is as follows. Section 1 is concerned principally with the problem formulation. The necessary and sufficient conditions for optimality are in the "subdifferential" form and do not require the differentiability of cost functional.

Thus our result is very much in the spirit of Rockafellar works [15, 16]. These conditions are spelled out in Theorem 1, whose proof is set forth in Section 3. In Section 4 we formulate a control problem for linear hereditary differential systems with convex criterion. The necessary and sufficient conditions are then specialised for this particular problem (see Theorem 2). These results may be compared with those of R. Datko [7, 8], H. T. Banks and M. Q. Jacobs [1], H. T. Banks and G. A. Kent [2], M. C. Delfour and S. K. Mitter [9], A. Halanay [11] (further references may be found in these papers).

However, our results on necessary optimality conditions in problems involving linear evolution equations in Hilbert space setting (see Lions's book [12] for significant results in this field) differ from previous results (even specialised to functional differential equations) which involve stronger regularity assumptions than ours. For the most part, these papers are concerned with quadratic cost criterion and certainly the methods used here are very different. Our approach is much similar to that used by the author in [3] and it relies on the methods of convex analysis. For significant results in this field, relevant to this paper we refer to the books of Rockafellar [13] and Brézis [6].

1. Notation and Formulation of the Control Problem

We first introduce some preliminary notations.

1. H will be a real Hilbert space identified with its own dual. The norm on H will be denoted by $|\cdot|$ and the inner product by $(,)$.

2. U will denote a real Hilbert space with inner product $\langle , \rangle$ and norm $\|\cdot\|$.

3. Let $[0, T]$ be a finite closed interval of real line. Then $L^2(0, T; H)$ will denote the usual Hilbert space of "square integrable" functions from $]0, T[$ to H. By $C(0, T; H)$ we shall denote the space of all continuous functions from $[0, T]$ to H endowed with the usual norm. Let $\mathscr{M}(0, T; H)$ be the dual space of $C(0, T; H)$. In other words, $\mathscr{M}(0, T; H)$ is the space of all bounded H-valued measures on $[0, T]$. The norm on $\mathscr{M}(0, T; H)$ is given by $\|\mu\|_{\mathscr{M}} = \sup\{|\mu(\varphi)|; \|\varphi\|_{C(0,T;H)} \leqslant 1\}$ and denote by $\mu(\varphi)$ the value of $\mu \in \mathscr{M}(0, T; H)$ at $\varphi \in C(0, T; H)$.

4. $BV(0, T; H)$ will denote the Banach space of all functions $x\colon [0, T] \to H$ which are of bounded variation. If x' denotes the derivative of $x \in BV(0, T; H)$ taken in the sense of vectorial distributions over $]0, T[$ then $x' \in \mathscr{M}(0, T; H)$ (cf. [6, Proposition A.5]).

5. Let K be a closed convex subset of H and let $\mathscr{K}$ be the subset of $C(0, T; H)$ defined by

$$\mathscr{K} = \{x \in C(0, T; H); x(t) \in K \text{ for } t \in [0, T]\}.$$

We denote by $\mathscr{N}(x, K) \subset \mathscr{M}(0, T; H)$ the cone of normals to $\mathscr{K}$ at x. This is the closed, convex cone defined by

$$\mathscr{N}(x, K) = \{\mu \in \mathscr{M}(0, T; H); \mu(x - y) \geqslant 0 \text{ for all } y \in \mathscr{K}\}.$$

The following functions and operators will be used to define the control problem (1.1), (1.2), and (1.3) below.

6. $S(t, s)$ will denote a family of linear continuous operators from H into itself which are strongly continuous in the triangle $\Delta = \{(t, s); 0 \leqslant s \leqslant t \leqslant T\}$ and satisfy the following properties:

(a) $S(t, s) = S(t, \tau)\, S(\tau, s) \qquad 0 \leqslant s \leqslant \tau \leqslant t \leqslant T.$

(b) $S(t, t) = I \qquad 0 \leqslant t \leqslant T.$

(c) $\|S(t, s)\|_{L(H,H)} \leqslant M \qquad \text{for } 0 \leqslant s \leqslant t \leqslant T.$

The adjoint system $S^*(t, s)\colon \Delta \to L(H, H)$ also is assumed to be strongly continuous.

7. The mapping $B(t)$: $]0, T[\to L(U, H)$ will be assumed to be strongly measurable as well as the adjoint mapping $B^*(t)$. Furthermore, $B \in L^\infty(0, T; L(U, H))$.

8. L and l are lower semicontinuous and convex functions defined on $H \times U$ and $H \times H$, respectively, with values in $]-\infty, +\infty]$, not identically $+\infty$.

The type of control problem which we shall examine is of the following type

$$\text{Minimize} \int_0^T L(x(t), u(t))\, dt + l(x(0), x(T)) \tag{1.1}$$

in $x \in C(0, T; H)$ and $u \in L^2(0, T; U)$, subject to the constraints

$$x(t) = S(t, 0)\, x(0) + \int_0^t S(t, s)\, B(s)\, u(s)\, ds, \qquad 0 \leqslant t \leqslant T \tag{1.2}$$

$$x(t) \in K \qquad \text{for every } t \in [0, T]. \tag{1.3}$$

If $S(t, s)$ is the ***evolution operator*** associated with a family $\{A(t); 0 \leqslant t \leqslant T\}$ of unbounded closed linear operators acting in H, then $x(t)$ given by variation constants formula (1.2) is just the weak solution ("mild" solution in other terminology) of time dependent evolution equation

$$(dx/dt)\,(t) = A(t)\, x(t) + B(t)\, u(t), \qquad 0 < t < T. \tag{1.4}$$

It should be emphasised that (1.1) includes as special cases various control problems associated with state equation (1.2). In fact, the end point constraint

$$x(0) \in X_0\,, \qquad x(T) \in X_T$$

where X_0 and X_T are closed convex subsets of H, can be incorporated into the problem by redefining

$$l(x_1\,, x_2) = +\infty \qquad \text{if} \qquad x_1 \bar{\in} X_0 \quad \text{or} \quad x_2 \bar{\in} X_T\,.$$

The reader is referred to [15] for further discussion and examples.

We call an end point pair $[x_0\,, x_T] \in H \times H$ ***attainable*** for L if there are $x \in C(0, T; H)$ and $u \in L^2(0, T; U)$ such that $x(0) = x_0$, $x(T) = x_T$ and

$$x(t) = S(t, 0)\, x(0) + \int_0^t S(t, s)\, B(s)\, u(s)\, ds \qquad \text{on} \qquad [0, T] \tag{1.5}$$

$$x(t) \in K \qquad \text{for } t \in [0, T] \tag{1.6}$$

$$L(x, u) \in L^1(0, T). \tag{1.7}$$

The set of all attainable pairs will be denoted by C_L. We shall also denote by $D(l)$ the effective domain of l, i.e.,

$$D(l) = \{(h_1, h_2) \in H \times H; l(h_1, h_2) < +\infty\}.$$

Let $H: H \times U \to [-\infty, +\infty]$ be the Hamiltonian function corresponding to L. In other words

$$H(x, p) = \sup\{\langle p, u\rangle - L(x, u); u \in U\}. \tag{1.8a}$$

We set

$$\mathrm{Dom}_1 H = \{x \in H; H(x, p) > -\infty \text{ for every } p \in U\},$$

$$\mathrm{Dom}_2 H = \{p \in U; H(x, p) < +\infty \text{ for every } x \in H\}.$$

The set $\mathrm{Dom}\, H = \mathrm{Dom}_1 H \times \mathrm{Dom}_2 H$ is called the effective domain of H.

For the study of problem (1.1)–(1.3) one needs further conditions on L, l and K. The first is:

(A) $\mathrm{Dom}_1 H = H$ *and* $0 \in \mathrm{int}\, \mathrm{Dom}_2 H$.

The condition that $\mathrm{Dom}_1 H = H$ implies that for every $x \in H$ there exists at least an element $u \in U$ such that $L(x, u) < +\infty$. The rest of the assumptions ($0 \in \mathrm{int}\, \mathrm{Dom}_2 H$) is a growth condition of Nagumo–Tonelli type on $L(x, v)$ as a function of v. In fact it can be equivalently expressed as:

$$\lim_{\|u\| \to +\infty} L(x, u)/\| u \| \geqslant \rho > 0 \qquad \text{for all } x \in H. \tag{1.8b}$$

(B) *There is at least one pair of functions* $(x, u) \in C(0, T; H) \times L^2(0, T; U)$ *satisfying* (1.5, 1.7) *and*

$$l(x(0), x(T)) < +\infty, \qquad x(t) \in \mathrm{int}\, K \quad \textit{for every } t \in [0, T]. \tag{1.9}$$

(C) *There is at least one attainable pair* $(x_0, x_T) \in C_L \cap D(l)$ *such that one of the following two conditions holds:*

$$x_T \in \mathrm{int}\{h \in H; (x_0, h) \in C_L\}. \tag{1.10}$$

$$x_T \in \mathrm{int}\{h \in H; (x_0, h) \in D(l)\}. \tag{1.11}$$

2. Necessary and Sufficient Conditions for Optimal Control

We assume familiarity with the definitions and basic results in theory of convex functions defined on infinite dimensional spaces. However, for easy reference we review some basic facts about conjugates and subdifferentials.

Given a lower semicontinuous convex function φ from a Banach space X to $]-\infty, +\infty]$ and a point $x \in X$, we denote by $\partial\varphi(x)$ the set of all $x^* \in X^*$ such that

$$\varphi(x) \leqslant \varphi(y) + (x - y, x^*) \qquad \text{for every } y \in X.$$

Here X^* is the dual space of X and (x, x^*) is the value of $x^* \in X^*$ at $x \in X$. Such vectors x^* are called ***subgradients*** of φ at x, and $\partial\varphi(x)$ is called the ***subdifferential*** of φ at x. For every $x \in X$, $\partial\varphi(x)$ is a closed convex subset (possibly empty) of X. If φ happens to be Gâteaux differentiable at x, then $\partial\varphi(x)$ consists of a single element, namely the gradient $\nabla\varphi(x)$ of φ at x.

The function $\varphi^*: X^* \to]-\infty, +\infty]$ defined by

$$\varphi^*(x^*) = \sup\{(x, x^*) - \varphi(x); x \in X\}$$

is called the ***conjugate*** of φ. The function φ^* is always convex and lower semicontinuous on X^*. In the case where X is reflexive one has $x^* \in \partial\varphi(x)$ if and only if $x \in \partial\varphi^*(x^*)$.

We shall denote by $\partial l(x_1, x_2)$ the subdifferential of the convex function l at (x_1, x_2). Thus $\partial l(x_1, x_2)$ is a certain closed convex subset of $H \times H$. The subdifferential $\partial L(x, u)$ of L at $(x, u) \in H \times U$ will be a closed convex subset of $H \times U$ which will be written (somewhat imperfectly) as

$$\partial L(x, u) = [\partial_x L(x, u), \partial_u L(x, u)]$$

where $\partial_x L(x, u) \subset H$ and $\partial_u L(x, u) \subset U$. If L is actually (finite and) Gâteaux differentiable at (x, u), then $\partial_x L(x, u) = \nabla_x L(x, u)$ and $\partial_u L(x, u) = \nabla_u L(x, u)$.

We shall say that a given pair $(x, u) \in C(0, T; H) \times L^2(0, T; U)$ is ***extremal*** for problem (1.1) $\sim$ (1.3) if there exist functions $q \in L^2(0, T; H)$, $w \in BV(0, T; H)$ and $p: [0, T] \to H$, bounded and strongly measurable on $[0, T]$ such that

$$x(t) = S(t, 0)\, x(0) + \int_0^t S(t, s)\, B(s)\, u(s)\, ds \qquad \text{on} \qquad [0, T] \tag{2.1}$$

$$p(t) = S^*(T, t)\, p(T) - \int_t^T S^*(s, t)\, q(s)\, ds - \int_t^T S^*(s, t)\, dw(s), \qquad 0 \leqslant t \leqslant T \tag{2.2}$$

$$w' \in \mathscr{N}(x, K) \tag{2.3}$$

$$\{B^*(t)\, p(t), q(t)\} \in \partial L(x(t), u(t)) \qquad \text{a.e.} \qquad t \in]0, T[\tag{2.4}$$

$$\{p(0), -p(T)\} \in \partial l(x(0), x(T)) \qquad \text{(transversality).} \tag{2.5}$$

As mentioned before the distributional derivative w' of $w \in BV(0, T; H)$ is a H-valued bounded measure μ on $[0, T]$. In this context, $\int_t^T S(s, t)\, dw(s)$

denotes the integral of $S^*(0, t)$ over $[t, T]$ with respect to measure μ generated by w on $[0, T]$. In order to avoid some technical discussions concerning the applicability of Fubini's theorem, we shall mean by $\int_t^T S^*(s, t)\, dw(s)$ the function $g(t)$: $[0, T] \to H$ defined by

$$\int_0^T (g(t), \psi(t))\, dt = \mu \left(\int_0^t S^*(t, s)\, \psi(s)\, ds \right) \qquad \text{for every } \psi \in L^1(0, T; H).$$

It should be noted that if $K = H$ (i.e., are no state constraints) then $\mathcal{N}(x, K) = \{0\}$ so that equation (2.2) shows that the dual extremal arc $p(t)$ is continuous in t on $[0, T]$.

To be more specific, let us suppose that $S(t, s) = S(t - s)$ is an one parameter semigroup and let A be its infinitesimal generator. Then the function $p(t) = -\int_t^T S^*(s - t)\, dw(s)$ satisfies in the sense of vectorial distributions over $]0, T[$ the equation

$$p' + A^* p = \mu.$$

Inasmuch as $\mu \in \mathcal{M}(0, T; H)$ we may infer that $p' \in \mathcal{M}(0, T; (D(A))^*)$ so that $p \in BV(0, T; (D(A))^*)$. Here $(D(A))^*$ denotes the dual space of $D(A)$ endowed with graph norm.

The main result of this paper may be stated as follows.

THEOREM 1. *Assume the existence of spaces, functionals and operators satisfying the above hypotheses. Then, in order that the pair* $[x, u]$ *be optimal in the problem* (1.1) $\sim$ (1.3), *it is necessary and sufficient that it is extremal.*

3. PROOF OF THEOREM 1

1. *Sufficiency.* Let $[x, u]$ be extremal. Fixing $(y, v) \in C(0, T; H) \times L^2(0, T; U)$, we observe from (2.4) and the definition of "subgradient" that

$$L(x,u) \leqslant L(y,v) + (q, x - y) + (p, B(t)(u - v)) \quad \text{a.e. on} \quad]0, T[. \tag{3.1}$$

If y and v satisfy Eq. (1.2) and (1.3) then by (2.2) we have

$$\int_0^T (p(t), B(t)(u(t) - v(t)))\, dt$$
$$= \int_0^T (p(T), S^*(T, t)\, B(t)(u(t) - v(t)))\, dt$$
$$- \int_0^T \left(B(t)(u(t) - v(t)), \int_t^T S^*(s, t)\, q(s)\, ds\right) dt - \int_0^T (B(t)(u(t) - v(t)),$$

$$\int_t^T S^*(s, t)\, dw(s))\, dt$$

$$= -(p(T), S(T, 0)\,(x(0) - y(0))) + (p(T), x(T) - y(T))$$

$$- \int_0^T (q(s), x(s) - y(s))\, ds - \left(x(0) - y(0), \int_0^T S^*(s, 0)\, q(s)\, ds\right)$$

$$- \int_0^T \left(B(t)\,(u(t) - v(t)), \int_t^T S^*(s, t)\, dw(s)\right) ds$$

and therefore

$$\int_0^T (p(t), B(t)\,(u(t) - v(t)))\, dt$$

$$= (p(T), x(T) - y(T)) - (p(0), x(0) - y(0))$$

$$- \left(x(0) - y(0), \int_0^T S^*(s, 0)\, dw(s)\right) - \int_0^T (q(s), x(s) - y(s))\, ds$$

$$- \int_0^T \left(B(t)\,(u(t) - v(t)), \int_t^T S^*(s, t)\, dw(s)\right) dt. \tag{3.2}$$

Here we have used in particular, the property of $S(s, t)$ and Fubini's theorem.

On the other hand, interchanging the order of integration, which is easily justified by Fubini's theorem, yields

$$\int_0^T \left(B(t)\,(u(t) - v(t)), \int_t^T S^*(s, t)\, dw(s)\right) dt$$

$$= \int_0^T \left(S^*(s, t)\, dw(s), \int_0^s B(t)\,(u(t) - v(t))\, dt\right)$$

$$= \mu(x(t) - y(t)) - \mu(S(t, 0)\,(x(0) - y(0))).$$

Hence

$$\left(x(0) - y(0), \int_0^T S^*(s, 0)\, dw(s)\right) + \int_0^T \left(B(t)\,(u(t) - v(t)), \int_t^T S^*(s, t)\, dw(s)\right) dt$$

$$= \mu(x - y) \geqslant 0,$$

because $\mu = w' \in \mathscr{N}(x, K)$. Comparing this inequality with (3.1) and (3.2) we conclude that

$$\int_0^T L(x, u)\, dt \leqslant \int_0^T L(y, v)\, dt + (p(T), x(T) - y(T)) - (p(0), x(0) - y(0))$$

while by (2.5) we have

$$(p(0), x(0) - y(0)) - (p(T), x(T) - y(T)) \geqslant l(x(0), x(T)) - l(y(0), y(T)).$$

Since y, v were arbitrary we may conclude that (x, u) is optimal for the problem (1.1)–(1.3).

2. *Necessity.* The proof of the necessity in Theorem 1 is more complicated so we shall divide it in several steps. We define

$$L_\lambda(x,l) = \inf \left\{ \frac{|x - y|^2 + \| u - v \|^2}{2\lambda} + L(y, v); y \in H, v \in U \right\}$$

$$l_\lambda(h_1, h_2) = \inf \left\{ \frac{|h_1 - \bar{h}_1|^2 + |h_2 - \bar{h}_2|^2}{2\lambda} + l(\bar{h}_1, \bar{h}_2); \bar{h}_1, \bar{h}_2 \in H \right\}$$

and

$$\varphi_\lambda(x) = \inf\{| x - y |^2/2\lambda; y \in K\}.$$

We recall (see e.g. [6]) that for every $\lambda > 0$ the functions L_λ, l_λ and φ_λ are convex, everywhere finite and Fréchet differentiable on $H \times U$, $H \times H$ and H, respectively.

Let $[x, u]$ be an optimal pair of the problem (1.1)–(1.3). The first step of the proof is given by Lemma 1 below.

LEMMA 1. *For every $\lambda > 0$ there exist $x_\lambda \in C(0, T; H)$, $u_\lambda \in L^2(0, T; H)$ and $p_\lambda \in C(0, T; H)$ such that*

$$x_\lambda(t) = S(t, 0)\, x_\lambda(0) + \int_0^t S(t, s)\, B(s)\, u_\lambda(s) \qquad \textit{for } 0 \leqslant t \leqslant T \tag{3.3}$$

$$p_\lambda(t) = S^*(T, t)\, p_\lambda(T) - \int_t^T S^*(s, t)\, \partial_x L_\lambda(x_\lambda, u_\lambda)\, ds - \int_t^T S^*(s, t)\, \partial\varphi_\lambda(x_\lambda)\, ds, \tag{3.4}$$

$$B^*(t)\, p_\lambda(t) + u(t) - u_\lambda(t) = \partial_u L_\lambda(x_\lambda(t), u_\lambda(t)) \qquad \text{a.e.} \qquad t \in]0, T[\tag{3.5}$$

$$[p_\lambda(0) + x(0) - x_\lambda(0), -p_\lambda(T)] = \partial l_\lambda(x_\lambda(0), x_\lambda(T)). \tag{3.6}$$

Furthermore,

$$\lim_{\lambda \to 0} x_\lambda = x \qquad \textit{in} \qquad C(0, T; H). \tag{3.7}$$

$$\lim_{\lambda \to 0} u_\lambda = u \qquad \textit{strongly in} \qquad L^2(0, T; U). \tag{3.8}$$

Proof. Let $F_\lambda: H \times L^2(0, T; U) \to]-\infty, +\infty]$ be the convex function defined by

$$F_\lambda(h, v) = \int_0^T (L_\lambda(y(t), v(t)) + \tfrac{1}{2} \| u(t) - v(t)\|^2 + \varphi_\lambda(y(t)) \, dt$$
$$+ l_\lambda(h, y(T)) + \tfrac{1}{2} | h - x(0)|^2$$

where

$$y(t) = S(t, 0) \, h + \int_0^t S(t, s) \, B(s) \, v(s) \, ds, \qquad 0 \leqslant t \leqslant T.$$

Obviously, F_λ attains its infimum on $H \times L^2(0, T; U)$ in a unique point (h_λ, u_λ). Setting

$$x_\lambda(t) = S(t, 0) \, h_\lambda + \int_0^t S(t, s) \, B(s) \, u_\lambda(s) \, ds$$

we may write

$$F_\lambda(h_\lambda, u_\lambda) = \int_0^T (L_\lambda(x_\lambda(t), u_\lambda(t)) + \tfrac{1}{2} \| u(t) - u_\lambda(t)\|^2 + \varphi_\lambda(x_\lambda(t)) \, dt$$
$$+ l_\lambda(x_\lambda(0), x_\lambda(T)) + \tfrac{1}{2} | x_\lambda(0) - x(0)|^2. \tag{3.9}$$

This implies by a standard argument that

$$\int_0^T [(q_\lambda(t) + \partial\varphi_\lambda(x_\lambda(t)), y(t)) + \langle \partial_u L(x_\lambda(t), u_\lambda(t)) + u_\lambda(t) - u(t), v(t)\rangle] \, dt$$
$$+ (h_\lambda^1 + x_\lambda(0) - x(0), y(0)) + (h_\lambda^2, y(T)) = 0 \tag{3.10}$$

for all $y \in C(0, T; H)$ and $v \in L^2(0, T; U)$ which satisfy

$$y(t) = S(t, 0) \, y(0) + \int_0^t S(t, s) \, B(s) \, v(s) \, ds, \qquad \text{on} \qquad [0, T]. \tag{3.11}$$

Here we have used the notations

$$q_\lambda = \partial_x L_\lambda(x_\lambda, u_\lambda); \qquad [h_\lambda^1, h_\lambda^2] = \partial l_\lambda(x_\lambda(0), x_\lambda(T)).$$

Let $p_\lambda \in C(0, T; H)$ be defined by

$$p_\lambda(t) = -S^*(T, t) \, h_\lambda^2 - \int_t^T S^*(s, t) \, (q_\lambda(s) + \partial\varphi_\lambda(x_\lambda(s))) \, ds, \qquad 0 \leqslant t \leqslant T.$$

Then

$$\int_0^T \langle B^* p_\lambda(t), v(t)\rangle \, dt = - \int_0^T \left(\left(q_\lambda(s) + \partial\varphi_\lambda(x_\lambda(s)), \int_0^s S(s, t) \, B(t) \, v(t)\right) dt\right) ds$$
$$- \int_0^T (h_\lambda^2, S(T, t) \, B(t) \, v(t)) \, dt$$

and a simple calculation involving (3.10) and (3.11), yields

$$B^*(t)\, p_\lambda(t) = \partial_u L_\lambda(x_\lambda(t), u_\lambda(t)) + u_\lambda(t) - u(t) \quad \text{a.e.} \quad \text{on} \quad]0, T[$$
$$p_\lambda(0) = h_\lambda^1 + x_\lambda(0) - x(0).$$

Summarizing to this point, we have shown that x_λ, u_λ and p_λ satisfy Eqs. (3.3), (3.4), (3.5) and (3.6).

Recalling that

$$F_\lambda(h_\lambda, u_\lambda) \leqslant F_\lambda(h, v) \qquad \text{for every } (h, v) \in H \times L^2(0, T; U)$$

it follows from (3.9) that

$$\begin{aligned} &\tfrac{1}{2}\int_0^T \| u_\lambda(t) - u(t)\|^2\, dt + \tfrac{1}{2}\, | x_\lambda(0) - x(0)|^2 \\ &\quad \leqslant \int_0^T L(x, u)\, dt - \int_0^T (L_\lambda(x_\lambda, u_\lambda) + \varphi_\lambda(x_\lambda))\, dt + l(x(0), x(T)) \\ &\quad - l_\lambda(x_\lambda(0), x_\lambda(T)) \end{aligned} \tag{3.12}$$

because $L_\lambda(x, u) \leqslant L(x, u)$, $\varphi_\lambda(x) \leqslant \varphi(x)$ and $l_\lambda(x(0), x(T)) \leqslant l(x(0), x(T))$ for all $\lambda > 0$.

Since the subdifferential ∂L is maximal monotone in $(H \times U) \times (H \times U)$, the operator $(I + \lambda\partial L)^{-1}$ is well defined and nonexpansive on $H \times U$ (see e.g. [6]). Moreover, by definition of L_λ and l_λ we have

$$L_\lambda(x_\lambda, u_\lambda) = \frac{\lambda}{2}\| \partial L_\lambda(x_\lambda, u_\lambda)\|^2_{H\times U} + L((I + \lambda\partial L)^{-1}(x_\lambda, u_\lambda)) \tag{3.13}$$

$$l_\lambda(x_\lambda(0), x_\lambda(T)) = \frac{\lambda}{2}\| \partial l_\lambda(x_\lambda(0), x_\lambda(T))\|^2_{H\times H} + l((I + \lambda\partial l)^{-1}(x_\lambda(0), x_\lambda(T)). \tag{3.14}$$

Here we have used the same symbol I to denote the identity operator in $H \times U$ and $H \times H$, respectively. In particular, (3.13) and (3.14) imply that

$$\lim_{\lambda\to 0}[(I + \lambda\partial L)^{-1}(x_\lambda, u_\lambda) - (x_\lambda, u_\lambda)] = 0 \qquad \text{strongly in } L^2(0, T; H) \times L^2(0, T; U),$$

$$\lim_{\lambda\to 0}[(I + \lambda\partial l)^{-1}(x_\lambda(0), x_\lambda(T)) - (x_\lambda(0), x_\lambda(T))] = 0 \qquad \text{strongly in } H \times H.$$

Since $[u_\lambda, x_\lambda(0)]$ remain in a bounded subset of $L^2(0, T; U) \times H$ we may assume that

$$\begin{aligned} u_\lambda &\to \tilde{u} &\quad &\text{weakly in} \quad L^2(0, T; U) \\ x_\lambda &\to \tilde{x} &\quad &\text{weakly in} \quad L^2(0, T; H) \end{aligned}$$

and

$$[x_\lambda(0), x_\lambda(T)] \to [x(0), x(T)] \qquad \text{weakly in} \qquad H \times H.$$

We also observe that

$$\tilde{x}(t) = S(t, 0)\, \tilde{x}(0) + \int_0^t S(t, s)\, B(s)\, \tilde{u}(s)\, ds \qquad \text{for } t \in [0, T].$$

Furthermore, since $\int_0^T \varphi_\lambda(x_\lambda(t))\, dt$ is bounded, from the definition of φ_λ we may infer that $\tilde{x}(t) \in K$ for every $t \in [0, T]$.

The function $(y, v) \to \int_0^T L(y, v)\, dt$ being convex and lower semicontinuous on $L^2(0, T; H) \times L^2(0, T; U)$ it is weakly lower semicontinuous too. Thus by (3.13) we may infer that

$$\liminf_{\lambda \to 0} \int_0^T L_\lambda(x_\lambda, u_\lambda)\, dt \geqslant \int_0^T L(x, u)\, dt.$$

Next by (3.14)

$$\liminf_{\lambda \to 0} l_\lambda(x_\lambda(0), x_\lambda(T)) \geqslant l(x(0), x(T)).$$

Comparing these inequalities with (3.12), we conclude that $\tilde{x} = x$, $\tilde{u} = u$ and

$$\lim_{\lambda \to 0} x_\lambda(0) = x(0) \qquad \text{strongly in} \qquad H,$$

$$\lim_{\lambda \to 0} u_\lambda = u \qquad \text{strongly in} \qquad L^2(0, T; U).$$

In particular this implies that $\lim_{\lambda \to 0} x_\lambda(t) = x(t)$ uniformly in t on $[0, T]$, as claimed.

LEMMA 2. *Let $\{p_\lambda\}$ be the family of functions that appear in Lemma* 1. *Then there is a positive constant C independent of λ such that*

$$| p_\lambda(T) | \leqslant C \qquad \textit{for all } \lambda > 0. \tag{3.15}$$

Proof. Since the proof is essentially the same as in the author's paper [3], it will thus only sketched. For every $(h^1, h^2) \in H \times H$ we denote

$$\phi(h^1, h^2) = \inf\{G(v);\, v \in L^2(0, T; U);\, y(t) \in K \text{ on } [0, T];\, y(0) = h^1, y(T) = h^2\}$$

where

$$G(v) = \int_0^T (L(y(t), v(t)) + \tfrac{1}{2} \| v(t) - u(t) \|^2)\, dt$$

and $y(t) = S(t)\, y(0) + \int_0^t S(t, s)\, B(s)\, v(s)\, ds$ on $[0, T]$. The function G is convex, lower semicontinuous on $L^2(0, T; U)$ and $G(v) \to +\infty$ as

$\| v \|_{L_2(0,T;U)} \to +\infty$. Also observe that $D(\phi) = C_L$ and for every $(h^1, h^2) \in C_L$ the infimum defining $\phi(h^1, h^2)$ is attained. In particular, this implies that ϕ is convex, lower semicontinuous and nowhere $-\infty$ on $H \times H$.

First, we suppose that condition (1.10) in Hypothesis (C) holds. Thus, there is $y \in C(0, T; H)$ and $\rho > 0$ such that $l(y(0), y(T)) < +\infty$ and

$$\phi(y(0), y(T) + \rho h) \leqslant C \qquad \text{for all } h \in H, \quad | h | = 1$$

where C is some positive constant. Let $(v, z) \in L^2(0, T; U) \times C(0, T; H)$ be such that $z(t) \in K$ on $[0, T]$ and

$$z(t) = S(t, 0)\, z(0) + \int_0^t S(t, s)\, B(s)\, v(s)\, ds, \qquad 0 \leqslant t \leqslant T,$$

$$z(0) = y(0), \qquad z(T) = y(T) + \rho h$$

and

$$\int_0^T (L(z(t), v(t)) + \tfrac{1}{2} \| v(t) - u(t)\|^2)\, dt = \phi(z(0), z(T)).$$

Since $\varphi_\lambda(z) = 0$, it follows from Eqs. (3.3), (3.4) and (3.5) that

$$\int_0^T (L_\lambda(x_\lambda , u_\lambda) + \varphi_\lambda(x_\lambda) + \tfrac{1}{2} \| u_\lambda - u \|^2)\, dt - \int_0^T (L_\lambda(z, v) + \tfrac{1}{2} \| v - u \|^2)\, dt$$
$$\leqslant (p_\lambda(T), x_\lambda(T) - z(T)) - (p_\lambda(0), x_\lambda(0) - y(0)).$$

Hence

$$\rho(p_\lambda(T), h) \leqslant C + \phi(y(0), y(T) + \rho h) + (p_\lambda(T), x_\lambda(T)) - (p_\lambda(0); x_\lambda(0) - y(0))$$
$$\leqslant C + \phi(y(0), y(T) + \rho h) + l_\lambda(y(0), y(T)) - l_\lambda(x_\lambda(0), x_\lambda(T)).$$

Since $l_\lambda(y(0), y(T)) \leqslant l(y(0), y(T))$ and $l_\lambda(x_\lambda(0), x_\lambda(T))$ are uniformly bounded from below, we may infer that

$$\rho(p_\lambda(T), h) \leqslant C \qquad \text{for all } \lambda > 0 \quad \text{and} \quad | h | = 1.$$

(We shall denote by C several positive constants independent of λ). The letter implies (3.15) as claimed. Next, we shall assume that condition (1.11) holds and choose $y \in C(0, T; H)$ such that $[y(0), y(T)] \in C_L \cap D(l)$ and

$$y(T) \in \operatorname{int}\{h \in H;\ [y(0), h] \in D(l)\}.$$

This implies that function $h \to l(y(0), h)$ is locally bounded at $h = y(T)$. In other words there exist $\rho > 0$ and $C > 0$ such that

$$l(y(0), y(T) + \rho h) \leqslant C \qquad \text{for all } h \in H, \quad | h | = 1. \tag{3.16}$$

Again using the transversality conditions (3.6) we get

$$\rho(p_\lambda(T), h) \leqslant (p_\lambda(0), x_\lambda(0) - y(0)) - (p_\lambda(T), x_\lambda(T) - y(T)) \\ - (x_\lambda(0), x_\lambda(0) - x(0)) - l_\lambda(x_\lambda(0), x_\lambda(T)) + l_\lambda(y(0), y(T) + \rho h). \tag{3.17}$$

But making use once again by (3.3) $\sim$ (3.5) we deduce after same calculations that $\{(p_\lambda(0), x_\lambda(0) - y(0)) - (p_\lambda(T), x_\lambda(T) - y(T))\}$ is bounded. Thus by the same device as before, (3.16) and (3.17) again imply (3.15).

LEMMA 3. *There is a positive constant C independent of such that*

$$| p_\lambda(t)| \leqslant C \qquad on \qquad [0, T] \tag{3.18}$$

$$\int_0^T | \partial\varphi_\lambda(x_\lambda(t))| \, dt \leqslant C \tag{3.19}$$

for all sufficiently small $\lambda > 0$.

Proof. According to Hypothesis (B) there are $x_0 \in C(0, T; H)$, $u_0 \in L^2(0, T; U)$ and $\rho > 0$ such that

$$x_0(t) = S(t, 0)\, x_0(0) + \int_0^t S(t, s)\, B(s)\, u_0(s)\, ds$$

and

$$l(x_0(0), x_0(T)) < +\infty, \qquad x_0(t) + \rho h \in K \quad \text{for } t \in [0, T], \quad | h | = 1.$$

On the other hand, from the definition of $\partial\varphi_\lambda(x_\lambda)$, we have

$$(\partial\varphi_\lambda(x_\lambda), x_\lambda - x_0 - \rho h) \geqslant \varphi_\lambda(x_\lambda) - \varphi_\lambda(x_0 + \rho h)$$

while

$$\varphi_\lambda(x_0(t) + \rho h) = 0 \qquad \text{for every } t \in [0, T], \quad | h | = 1.$$

Hence

$$\rho \int_0^T | \partial\varphi_\lambda(x_\lambda(t))| \, dt \leqslant \int_0^T (\partial\varphi_\lambda(x_\lambda(t)), x_\lambda(t) - x_0(t))\, dt \qquad \text{for all } \lambda > 0. \tag{3.20}$$

Making use of Eq. (3.3), we get

$$\int_0^T (\partial_x L_\lambda(x_\lambda, u_\lambda) + \partial\varphi_\lambda(x_\lambda), x_\lambda - x_0)\, dt \\ = \Big(x_\lambda(0) - x_0(0), \int_0^T S^*(t, 0)\, (\partial_x L_\lambda(x_\lambda, u_\lambda) + \partial\varphi_\lambda(x_\lambda))\, dt \\ + \int_0^T (\partial\varphi_\lambda(x_\lambda(t)) + \partial_x L_\lambda(x_\lambda(t), u_\lambda(t)), \int_0^t S(t, s)\, B(s)\, (u_\lambda(s) - u_0(s))\, ds\Big)\, dt. \tag{3.21}$$

Next, we apply Fubini's theorem and substitute the right side of (3.4) into (3.21) to obtain,

$$
\begin{aligned}
\int_0^T & (\partial_x L_\lambda(x_\lambda, u_\lambda) + \partial\varphi_\lambda(x_\lambda), x_\lambda - x_0)\, dt \\
&= -(x_\lambda(0) - x_0(0), p_\lambda(0) - S^*(T,0)\, p_\lambda(T)) \\
&\quad - \int_0^T \langle \partial_u L_\lambda(x_\lambda, u_\lambda) + u_\lambda - u, u_\lambda - u_0\rangle\, dt \qquad (3.22) \\
&\quad + (p_\lambda(T), x_\lambda(0) - x_0(0)) - S(T,0)\,(x_\lambda(0) - x_0(0)) \\
&= -(p_\lambda(0), x_\lambda(0) - x_0(0)) + (p_\lambda(T), x_\lambda(T) - x_0(T)) \\
&\quad - \int_0^T \langle \partial_u L(x_\lambda, u_\lambda) + u_\lambda - u, u_\lambda - u_0\rangle\, dt.
\end{aligned}
$$

We observe from (3.6, 3.22) and the definition of subgradient that

$$
\begin{aligned}
\int_0^T & ((\partial_x L_\lambda(x_\lambda, u_\lambda), x_\lambda - x_0) + \langle \partial_u L_\lambda(x_\lambda, u_\lambda) + u_\lambda - u, u_\lambda - u_0\rangle)\, dt \\
&\quad - (p_\lambda(T), x_\lambda(T) - x_0(T)) + (p_\lambda(0), x_\lambda(0) - x_0(0)) \\
&\geqslant \int_0^T (L_\lambda(x_\lambda, u_\lambda) + \tfrac{1}{2}\| u_\lambda - u\|^2)\, dt - \int_0^T (L_\lambda(x_0, u_0) + \tfrac{1}{2}\| u - u_0\|^2)\, dt \\
&\quad + l_\lambda(x_\lambda(0), x_\lambda(T)) - l_\lambda(x_0(0), x_0(T)) + (x_\lambda(0) - x(0), x_\lambda(0) - x_0(0)).
\end{aligned}
$$

Since $L_\lambda(x_0, u_0) \leqslant L(x_0, u_0) \in L^1(0, T)$, $l_\lambda(x_0(0), x_0(T)) \leqslant l(x_0(0), x_0(T)) \leqslant +\infty$, it follows from (3.22) that

$$\int_0^T (\partial\varphi_\lambda(x_\lambda), x_\lambda - x_0)\, dt \leqslant C \qquad \text{for all } \lambda > 0$$

because $\{x_\lambda, u_\lambda\}$ is bounded in $C(0, T; H) \times L^2(0, T; U)$ and L_λ, l_λ are uniformly bounded from below by affine functions. Thus, comparing this inequality with (3.20) we obtain the estimate (3.19) as desired. In order to prove (3.18) we shall use Hypothesis (A). Let $\partial H = \{-\partial_x H, \partial_p H\}$ be the *subdifferential* of Hamiltonian function $H(x, p)$. In other words,

$$
\begin{aligned}
\partial_p H(x, p) &= \{v \in U;\, H(x, p) \leqslant H(x, \bar{p}) + \langle p - \bar{p}, v\rangle;\, \forall \bar{p} \in U\}, \\
\partial_x H(x, p) &= \{y \in H;\, H(x, p) \geqslant H(\bar{x}, p) + (x - \bar{x}, y);\, \forall x \in H\}.
\end{aligned}
$$

We note that $H(x, p)$ and $\partial H(x, p)$ are locally bounded at every point $(x, p) \in \operatorname{int} \operatorname{Dom} H$ (see R. T. Rockafellar [13, 14]). Thus, by our assumption, there is $\rho > 0$ such that

$$| H(x(t) + \rho h, 0)| \leqslant C \qquad \text{for } t \in [0, T], \quad | h | = 1, \qquad (3.23)$$

$$\sup\{\| z \|;\, z \in \partial_p H(x(t) + \rho h, 0)\} \leqslant C \qquad \text{for } t \in [0, T], \quad | h | = 1. \qquad (3.24)$$

(The function $x(t)$ being continuous, its range is compact in H). Let v_h: $[0, T] \to U$ be such that $v_h(t) \in \partial_p H(x(t) + \rho h, 0)$. Thus in virtue of the conjugacy correspondence between L and H, one has

$$L(x(t) + \rho h, v_h(t)) = -H(x(t) + \rho h, 0)$$

so that (3.23) and (3.24) imply that

$$L(x(t) + \rho h, v_h(t)) \leqslant C \qquad \text{for } t \in [0, T] \quad \text{and} \quad |h| = 1 \tag{3.25}$$

$$\| v_h(t)\| \leqslant C \qquad \text{for } t \in [0, T] \quad \text{and} \quad |h| = 1. \tag{3.26}$$

We set $q_\lambda(t) = \partial_x L_\lambda(x_\lambda(t), u_\lambda(t))$ and use once again the definition of ∂L_λ and Eq. (3.5) to obtain

$$\begin{aligned}(q_\lambda(t), x_\lambda(t) - x(t) - \rho h) + \langle B^*(t)\, p_\lambda(t) + u(t) - u_\lambda(t), u_\lambda(t) - v_h(t)\rangle \\ \geqslant L_\lambda(x_\lambda(t), u_\lambda(t)) - L_\lambda(x(t) + \rho h, v_h(t)), \quad \text{a.e.} \quad t \in]0, T[.\end{aligned}$$

In this inequality we take $h = q_\lambda(t)/| q_\lambda(t)|$ and use the estimate (3.25) to get

$$| q_\lambda(t)| \leqslant C_1 + C_2(| p_\lambda(t)| + \| u(t) - u_\lambda(t)\|) \| u_\lambda(t) - v_h(t)\| , \qquad \text{a.e.} \quad t \in]0, T[\tag{3.27}$$

for all sufficiently small λ. Here we have also used the fact that $\lim_{\lambda\to 0} x_\lambda(t) = x(t)$ uniformly in t on $[0, T]$.

Substitution of (3.15), (3.19) and (3.27) into (3.4) yields

$$\begin{aligned}| p_\lambda(t)| \leqslant C \left(1 + \int_t^T | p_\lambda(s) | \| u_\lambda(s) - v_h(s)\|\right) ds \\ + \int_t^T \| u(s) - u_\lambda(s)\| \| u_\lambda(s) - v_h(s)\| \, ds \qquad \text{for every } t \in [0, T].\end{aligned}$$

Using (3.25), the fact that $\{u_\lambda\}$ is bounded in $L^2(0, T; U)$ and Gronwall's inequality we may conclude that $\{| p_\lambda(t)|\}$ are uniformly bounded in t on $[0, T]$. This concludes the proof of Lemma 3.

We now turn to the proof of necessity in Theorem 1. First we observe that Lemma 1 and the estimate (3.27) imply that

$$\int_{E_\lambda} | q_\lambda(t)|^2 \, dt \leqslant C$$

where $\{E_\lambda\}$ is a family of measurable subsets of $[0, T]$ having the property that

$$\text{meas}\{]0, T[\setminus E_\lambda\} \to 0 \qquad \text{as} \qquad \lambda \to 0.$$

(In all cases where the notion of measure intervenes, Lebesgue measure is understood unless stated expressly to the contrary). Define

$$\tilde{q}_\lambda(t) = \begin{cases} q_\lambda(t) & \text{for } t \in E_\lambda , \\ 0 & \text{for } t \in [0, T]\backslash E_\lambda . \end{cases}$$

Then

$$\int_0^T | \tilde{q}_\lambda(t)|^2 \, dt \leqslant C, \tag{3.28}$$

for all sufficiently small $\lambda > 0$.

Now we may take weakly convergent subsequences. There exists a subsequence (again denoted λ) convergent to zero such that

$$\begin{array}{lcl} \tilde{q}_\lambda \to q & \text{weakly in} & L^2(0, T; H) \\ p_\lambda \to p & \text{weakly-star in} & L^\infty(0, T; H) \\ p_\lambda(0) \to p_0 & \text{weakly in} & H \\ p_\lambda(T) \to p_T & \text{weakly in} & H. \end{array}$$

Fixing (h_1 , h_2) in $H \times H$ we observe from (3.6) and the definition of ∂l_λ that

$$\begin{aligned} &(p_\lambda(0) + x(0) - x_\lambda(0), x_\lambda(0) - h_1) - (p_\lambda(T), x_\lambda(T) - h_2) \\ &\quad \geqslant l_\lambda(x_\lambda(0), x_\lambda(T)) - l_\lambda(h_1 , h_2) \end{aligned}$$

while

$$\lim_{\lambda \to 0} l_\lambda(h_1 , h_2) = l(h_1 , h_2), \qquad \liminf_{\lambda \to 0} l_\lambda(x_\lambda(0), x_\lambda(T)) \geqslant l(x(0), x(T)).$$

(This may be seen by the same device as that used in the proof of Lemma 1).

Therefore,

$$(p_0 , x(0) - h_1) - (p_T , x(T) - h_2) \geqslant l(x(0), x(T)) - l(h_1 , h_2)$$

because $[x_\lambda(0), x_\lambda(T)] \to [x(0), x(T)]$ strongly in $H \times H$. Thus we have shown that

$$[p_0 , -p_T] \in \partial l(x(0), x(T)). \tag{3.29}$$

Now we fix (y, v) in $L^2(0, T; H) \times L^2(0, T; U)$. We have

$$\begin{aligned} &(q_\lambda(t), x_\lambda(t) - y(t)) + \langle B^*(t) \, p_\lambda(t) + u(t) - u_\lambda(t), u_\lambda(t) - v(t)\rangle \\ &\quad \geqslant L_\lambda(x_\lambda(t), u_\lambda(t)) - L_\lambda(y(t), v(t)) \quad \text{a.e.} \quad \text{on} \quad]0, T[. \end{aligned} \tag{3.30}$$

Integrating both sides of (3.30) and letting $\lambda \to 0$, it follows from (3.28) that

$$\int_0^T ((q, x - y) + \langle B^*(t) p, u - v\rangle) \, dt \geqslant \int_0^T L(x, u) \, dt - \int_0^T L(y, v) \, dt \tag{3.31}$$

because $\lim_{\lambda \to 0} L_\lambda(y, v) = L(y, v)$ and $\lim_{\lambda \to 0} \inf \int_0^T L_\lambda(x_\lambda, u_\lambda) \, dt \geqslant \int_0^T L(x, u) \, dt$ (see the proof of Lemma 1). Since y and v were arbitrary, the integral inequality (3.31) implies the pointwise inequality

$$(q(t), x(t) - y) + \langle B^*(t) p(t), u(t) - v\rangle \geqslant L(x(t), u(t)) - L(y, v) \quad \text{a.e. on} \quad]0, T[$$

for all y in H and v in U. In other words, we have shown that

$$[q(t), B^*(t) p(t)] \in \partial L(x(t), u(t)) \quad \text{a.e.} \quad \text{on} \quad]0, T[. \tag{3.32}$$

Next, define

$$w_\lambda(t) = \int_0^t \partial\varphi_\lambda(x_\lambda(t)) \, dt, \qquad 0 \leqslant t \leqslant T.$$

Lemma 3 shows that $\{| w_\lambda(t)|\}$ is uniformly bounded in t on $[0, T]$. Then extracting further subsequence if necessary, we may assume that

$$w_\lambda \to w \qquad \text{weakly-star in} \qquad L^\infty(0, T; H)$$

as $\lambda \to 0$. Moreover, we have

$$\left| \int_0^T (w(t), \varphi'(t)) \, dt \right| \leqslant C \sup\{| \varphi(t))|; 0 \leqslant t \leqslant T\},$$

for every $\varphi \in \mathscr{D}(0, T; H)$. (We have denoted by $\mathscr{D}(0, T; H)$ the space of all infinitely differentiable functions φ: $[0, T] \to H$ with compact support in $]0, T[$). Thus we may conclude that $w(t)$ coincides almost everywhere on $]0, T[$ with a H-valued function (again denoted w) of bounded variation on $[0, T]$ (see [6, Proposition A.5]. Furthermore, the derivative w' of w taken in the sense of H-valued distributions over $]0, T[$ is a measure $\mu \in \mathscr{M}(0, T; H)$. From the definition of φ_λ we see that

$$\int_0^T (w_\lambda'(t), x_\lambda(t) - y(t)) \, dt \geqslant 0 \qquad \text{for every } y \in \mathscr{K}$$

where $\mathscr{K} = \{y \in C(0, T; H); y(t) \in K \text{ for } t \in [0, T]\}$. Since $w_\lambda' \to w'$ in the weak-star topology of $\mathscr{M}(0, T; H)$ we may infer that

$$\mu(x - y) \geqslant 0, \qquad \text{for } y \in \mathscr{K}.$$

In other words

$$\mu \in \mathcal{N}(x, K). \tag{3.33}$$

Let $\tilde{S}^*(t, s)$: $[0, T] \times [0, T] \to L(H, H)$ be defined by

$$\tilde{S}^*(t, s) = \begin{cases} S^*(t, s) & \text{for } 0 \leqslant s \leqslant t \leqslant T \\ I & \text{for } 0 \leqslant t \leqslant s \leqslant T. \end{cases}$$

Then

$$\int_t^T S^*(t, s)\, \partial\varphi_\lambda(x_\lambda(s))\, ds = \int_0^T \tilde{S}^*(t, s)\, w_\lambda'(s)\, ds - w_\lambda(t).$$

This shows that for every $t \in [0, T]$, $\lim_{\lambda\to 0} (p_\lambda(t) - w_\lambda(t)) = p(t) - w(t)$ exists in the weak topology of H and

$$p(t) = S^*(T, t)\, p_T - \int_t^T S^*(s, t)\, q(s)\, ds - \tilde{\mu}(\tilde{S}^*(\cdot, t)) + w(t) \qquad \text{for } t \in [0, T] \tag{3.34}$$

where $\tilde{\mu}$ is the vectorial measure defined by

$$(\tilde{\mu}(P), h) = \mu(Ph) \qquad \text{for every } h \in H \quad \text{and} \quad P \in L(H, H).$$

It should be observed by Fubini's theorem that $g(t) = -\tilde{\mu}(\tilde{S}^*(\cdot, t)) + w(t)$ may be equivalently defined as

$$\int_0^T (g(t), \psi(t))\, dt = \mu\left(\int_0^t S^*(t, s)\, \psi(s)\, ds\right), \qquad \text{for every } \psi \in L^1(0, T; H).$$

Finally, we note that $p_0 = p(0)$ and $p_T = p(T)$.

Then comparing (3.34) with (3.29), (3.32) and (3.33) we may infer that functions x, u, p and w satisfy the optimality conditions (2.1) $\sim$ (2.5). In other words, we have shown that (x, u) is extremal for the problem (1.1) $\sim$ (1.2) which completes the proof of Theorem 1.

4. Control of Linear Hereditary Differential Systems

We begin by reviewing some basic facts about linear *hereditary* differential systems (see [10] for a fuller exposition on the subject). The framework adopted in this paper and the general results stated below are essentially due to Delfour and Mitter [9].

Let $N \geqslant 1$ be an integer, let $a > 0$, $0 = \theta_0 > \theta_1, \ldots > \theta_N = -a$ be real numbers and $b \geqslant a$.

The symbol $\mathcal{L}_{p,q}$ denotes the space of all $p \times q$ real matrices endowed with a suitable norm. The symbol $\mathscr{L}^2(-b, 0; R^n)$ denotes the space of all "square integrable" functions endowed with the seminorm

$$\| y \|_{M^2}^2 = | y(0)|^2 + \int_{-b}^{0} | y(\theta)|^2 \, d\theta.$$

The quotient space of $\mathscr{L}^2(-b, 0; R^n)$ by the linear subspace of all y such that $\| y \|_{M^2} = 0$, will be denoted by $M^2(-b, 0; R^n)$, and its norm by $\| \cdot \|_{M^2}$. The space $M^2(-b; 0; R^n)$ is isometrically isomorphic to the product space $R^n \times L^2(-b, 0; R^n)$. In this context any $y \in M^2(-b, 0; R^n)$ will be expressed as $y = \{y^0, y^1\}$ where $y^0 \in R^n$ and $y^1 \in L^2(-b, 0; R^n)$. The space $M^2(-b, 0; R^n)$ is a Hilbert space with inner product

$$(x, y) = \langle x^0, y^0 \rangle + \int_{-b}^{0} \langle x^1(\theta), y^1(\theta) \rangle \, d\theta$$

for $x, y \in M^2(-b, 0; R^n)$ (the angle brackets on the right stand for the usual inner product in R^n). We shall also need $H^1(0, T; R^n)$, the Hilbert space of all absolutely continuous functions x: $[0, T] \to R^n$ with the property that the function $t \to x'(t) = (dx/dt)(t)$ belongs to $L^2(0, T; R^n)$. Finally, we shall use the symbol $W(-b, T; R^n)$ to denote the set of all functions $x(t)$: $[-b, T] \to R^n$ which beong to $H^1(0, T; R^n)$ on $[0, T]$ and are "square integrable" on $]-b, 0[$.

For $t \in [0, T]$, x_t denotes the function on $]-b, 0[$ defined by

$$x_t(\theta) = x(t + \theta), \qquad -b < \theta < 0.$$

Consider the affine hereditary differential system defined on $[0, T]$; $b < T < +\infty$,

$$\begin{aligned} x'(t) = A_{00}(t)\, x(t) + \sum_{i=1}^{N} A_i(t)\, x_t(\theta_i) + \int_{-b}^{0} A_{01}(t, \theta)\, x_t(\theta)\, d\theta \\ + B(t)\, u(t), \quad \text{a.e.} \quad \text{on} \quad]0, T[\end{aligned} \tag{4.1}$$

$$x_0 = h, \qquad h \in M^2(-b, 0; R^n) \tag{4.2}$$

where A_{00} and A_i $(i = 1, 2,..., N)$ are in $L^\infty(0, T; \mathscr{L}_{nn})$, $A_{01} \in L_\infty(0, T; -b, 0; \mathscr{L}_{nn})$, $B \in L^\infty(0, T; \mathscr{L}_{nm})$ and $u \in L^2(0, T; R^m)$. It should be said that under the above hypotheses (4.1) has a unique solution $x(\cdot, h, u)$ in $W(-b, T; R^n)$.

Let $\phi_0(t, s)$: $\Delta \times \Delta \to \mathscr{L}_{nn}$ be the unique solution in $H^1(s, T; \mathscr{L}_{nn})$ of the system

$$\begin{aligned} \frac{\partial \phi_0(t, s)}{\partial t} = A_{00}(t)\, \phi_0(t, s) + \sum_{i=1}^{N} A_i(t) \begin{Bmatrix} \phi_0(t + \theta_i, s), & t + \theta_i \geqslant s \\ 0, & \text{otherwise} \end{Bmatrix} \\ + \int_{-b}^{0} A_{01}(t, \theta) \begin{Bmatrix} \phi_0(t + \theta, s), & t + \theta \geqslant s \\ 0, & \text{otherwise} \end{Bmatrix} d\theta, \quad \text{a.e.} \quad \text{in} \quad]s, T[\end{aligned} \tag{4.3}$$

and let $\phi_1(t, s, \eta)\colon \Delta \times \Delta \times]-b, 0[\to \mathscr{L}_{nn}$, be defined by

$$\phi_1(t, s, \eta) = \sum_{i=1}^{N} \begin{cases} \phi_0(t, s+\eta-\theta_i)\, A_i(s+\eta-\theta_i), & \eta+s-t < \theta_i \leqslant \eta \\ 0, & \text{otherwise} \end{cases}$$
$$+ \begin{cases} \int_{-b}^{\eta} \phi_0(t, s+\eta-\theta)\, A_{01}(s+\eta-\theta, \theta)\, d\theta, & s+\eta \leqslant t-b \\ \int_{-t+s}^{\eta} \phi_0(t, s+\eta-\theta)\, A_{01}(s+\eta-\theta, \theta)\, d\theta, & s+\eta > t-b \end{cases}. \tag{4.4}$$

Let $S(t, s)\colon \Delta = \{(t, s)\colon 0 \leqslant s \leqslant t \leqslant T\} \to \mathscr{L}(M^2(-b, 0; R^n), M^2(-b, 0; R^n))$ be defined by

$$S(t, s)\, h = S_0(t, s)\, h^0 + S_1(t, s)\, h^1 \tag{4.5}$$

where $S_0(t, s) \in \mathscr{L}(R^n, M^2(-b, 0; R^n))$ and $S_1(t, s) \in \mathscr{L}(L^2(-b; R^n); M^2(-b, 0; R^n))$ are given by

$$[S_0(t, s)\, h^0]\,(\theta) = \begin{cases} \phi_0(t+\theta, s)\, h^0, & t+\theta \geqslant s \\ 0, & t+\theta < s \end{cases} \tag{4.6}$$

and

$$[S_1(t, s)\, h^1]\,(\theta) = \begin{cases} \int_{-b}^{0} \phi_1(t+\theta, s, \eta)\, h^1(\eta)\, d\eta, & t+\theta \geqslant s \\ h^1(t+\theta-s), & t+\theta < s \end{cases}. \tag{4.7}$$

It turns out that $S(t, s)$ satisfies the following properties

$$S(t, r) = S(t, s)\, S(s, r), \qquad T \geqslant t \geqslant s \geqslant r > 0, \quad S(t, t) = I, \quad 0 \leqslant t \leqslant T.$$

Define the operator $\tilde{B}(t)\colon R^n \to M^2(-b, 0; R^n)$ as

$$[\tilde{B}(t)\, u]\,(\theta) = \begin{cases} B(t)\, u, & \text{for } \theta = 0 \\ 0, & \text{otherwise.} \end{cases}$$

Let $\tilde{x}(t) = x_t\colon [0, T] \to M^2(-b, 0; R^n)$ be the unique solution of the system (4.1). Then $\tilde{x}$ can be expressed as

$$\tilde{x}(t) = S(t, 0)\, \tilde{x}(0) + \int_0^t S(t, s)\, \tilde{B}(s)\, u(s)\, ds, \qquad 0 \leqslant t \leqslant T. \tag{4.8}$$

Let $L\colon R^n \times R_m \to]-\infty, +\infty]$ and $l\colon M^2(-b, 0; R^n) \times M^2(-b, 0; R^n) \to]-\infty, +\infty]$ be given lower semicontinuous convex functions.

The control problem we shall examine is:

$$\text{Minimize} \int_0^T L(x(t), u(t))\, dt + l(x_0, x_T) \tag{4.9}$$

in $x \in W(-b, T; R^n)$ and $u \in L^2(0, T; R^m)$, subject to

$$x'(t) = A_{00}(t)\, x(t) + \sum_{i=1}^{N} A_i(t)\, x_t(\theta_i) + \int_{-b}^{0} A_{01}(t, \theta)\, x_t(\theta)\, d\theta$$
$$+ B(t)\, u(t) \quad \text{a.e.} \qquad t \in]0, T[\tag{4.10}$$

$$x(t) \in K \qquad \text{for every } t \in [0, T] \tag{4.11}$$

where K is a closed convex subset of R^n. It should be recalled that in (4.9) $x_0 \in M^2(-b, 0; R^n)$ represents the initial condition $[x(0), h^1(\theta)]$ into Eq. (4.10), while $x_T = [x(T), x(T+\theta)]$.

In particular, the optimization problem

$$\text{Minimize} \int_0^T L(x(t), u(t))\, dt$$

over all $(x, u) \in W(-b, T; R^n) \times L^2(0, T; R^m)$ satisfying (4.10), (4.11) and

$$x_0 \in X_0\,, \qquad x_T \in X_T$$

can be formulated as a problem of the type (4.9), setting

$$l(y_1, y_2) = \begin{cases} 0 & \text{if} \quad y_1 \in X_0 \quad \text{and} \quad y_2 \in X_T \\ +\infty & \text{otherwise.} \end{cases}$$

Here X_0 and X_T are nonempty closed convex subsets of $M^2(-b, 0; R^n)$. (In particular, X_0 or X_T may consist of a single point or be all of $M^2(-b, 0; R^n)$). As mentioned in Section 1, also there are implicit control constraint

$$u(t) \in U_0 = \{u \in R^m; L(x, u) < +\infty\}.$$

Let $\tilde{L}$: $M^2(-b, 0; R^n) \times R^m \to]-\infty, +\infty]$ be defined as

$$\tilde{L}(h, u) = L(h^0, u)$$

for $u \in R^m$ and $h = \{h^0, h^1\} \in R^n \times L^2(-b, 0; R^n) = M^2(-b, 0; R^n)$. In terms of $\tilde{x}$, $\tilde{B}$ and $\tilde{L}$ defined above, we can express problem (4.9) $\sim$ (4.11) as:

$$\text{Minimize} \int_0^T \tilde{L}(\tilde{x}(t), u(t))\, dt + l(x(0), x(T)) \tag{4.12}$$

in $\tilde{x} \in C(0, T; M^2(-b, 0; R^n))$ and $u \in L^2(0, T; R^m)$, subject to

$$\tilde{x}(t) = S(t, 0)\, \tilde{x}(0) + \int_0^t S(t, s)\, \tilde{B}(s)\, u(s)\, ds, \qquad 0 \leq t \leq T \tag{4.13}$$

$$\tilde{x}(t) \in \tilde{K} \qquad \text{for } t \in [0, T], \tag{4.14}$$

where

$$\tilde{K} = \{h = (h^0, h^1) \in M^2(-b, 0; R^n); h^0 \in K\}. \tag{4.15}$$

We are in the situation of general problem (1.1) $\sim$ (1.2) described in Section 1, where $H = M^2(-b, 0; R^n)$, $U = R^m$, $K = \tilde{K}$, $L = \tilde{L}$ and $S(t, s)$ defined as above. Extremality conditions (2.2)–(2.5) can be expressed in this case as

$$\tilde{p}(t) = S^*(T, t)\,\tilde{p}(T) - \int_t^T S^*(s, t)\,\tilde{q}(s)\,ds - \int_t^T S^*(s, t)\,d\tilde{w}(s), \qquad 0 \leqslant t \leqslant T \tag{4.16}$$

$$\tilde{w}' \in \mathcal{N}(\tilde{x}, \tilde{K}) \tag{4.17}$$

$$\tilde{B}^*(t)\,\tilde{p}(t) \in \partial_u \tilde{L}(\tilde{x}(t), u(t)); \qquad \tilde{q}(t) \in \partial_x \tilde{L}(\tilde{x}(t), u(t)) \quad \text{a.e.} \quad \text{on} \quad]0, T[\tag{4.18}$$

$$\{\tilde{p}(0), -\tilde{p}(T)\} \in \partial l(\tilde{x}(0), \tilde{x}(T)) \tag{4.19}$$

where $\tilde{p}(t) = [p(t), p_1(t, \theta)]: [0, T] \to R^n \times L^2(-b, 0; R^n)$ and $\tilde{w} = [w, w_1] \in BV(0, T; R^n) \times BV(0, T; L^2(-b, 0; R^n))$.

Observe that $\partial\tilde{L} = \{\partial L, 0\}$ and $\mathcal{N}(\tilde{x}, \tilde{K}) = \{\mathcal{N}_0(x, K), 0\} \subset \mathcal{M}(0, T; R^n) \times \mathcal{M}(0, T; L^2(-b, 0; R^n))$, where

$$\mathcal{N}_0(x, K) = \{\mu \in \mathcal{M}(0, T; R^n); \mu(x - y) \geqslant 0 \text{ for } y \in C(0, T; R^n), \\ y(t) \in K \text{ on } [0, T]\}. \tag{4.20}$$

Thus $\tilde{q}(t) = [q(t), 0]$ and $\tilde{w}(t) = [w(t), 0]$, where

$$q(t) \in \partial_x L(x(t), u(t)) \quad \text{a.e.} \quad \text{on} \quad]0, T[, \quad q \in L^2(0, T; R^n) \tag{4.21}$$

and

$$w \in BV(0, T; R^n), \qquad w' = \mu \in \mathcal{N}_0(x, K). \tag{4.22}$$

On the other hand, it follows from (4.5), (4.6), and (4.7) that

$$S^*(t, s)\,h = [S_0^*(t, s)\,h, S_1^*(t, s)\,h] \qquad \text{for } h \in M^2(-b, 0; R^n) \tag{4.23}$$

where

$$S_0^*(t, s)\,h = \phi_0^*(t, s)\,h^0 + \int_{s-t}^0 \phi_0^*(t + \theta, s)\,h^1(\theta)\,d\theta, \qquad 0 \leqslant s \leqslant t \leqslant T \tag{4.24}$$

and

$$(S_1^*(t, s)\,h)(\eta) = \phi_1^*(t, s, \eta)\,h^0 + \int_{-b}^{s-t} \phi_1^*(t + \theta, s, \eta)\,h^1(\theta)\,d\theta \\ + \begin{cases} h^1(\eta + s - t), & \eta > t - b - s \\ 0, & \eta < t - b - s \end{cases}, \tag{4.25}$$

$$\text{for } 0 \leqslant s \leqslant t \leqslant T, \quad \eta \in]-b, 0[.$$

Then a simple calculation involving (4.16), (4.23), (4.24) and (4.25) yields

$$p(t) = \phi_0{}^*(T, t)\, p(T) + \int_{t-T}^{0} \phi_0{}^*(T - \eta, t)\, p_1(T, \eta)\, d\eta$$
$$- \int_t^T \phi_0{}^*(s, t)\, q(s)\, ds - \int_t^T \phi_0{}^*(s, t)\, dw(s) \qquad \text{on} \qquad [0, T]. \quad (4.26)$$

$$p_1(t, \theta) = \phi_1{}^*(T, t, \theta)\, p(T) + \int_{-b}^{t-T} \phi_1{}^*(T + \eta, t, \theta)\, p_1(T, \eta)\, d\eta$$
$$+ \begin{cases} p_1(T, \theta + t - T), & \theta > T - b - t \\ 0, & \theta < T - b - t \end{cases} - \int_t^T \phi_1{}^*(s, t, \theta)\, q(s)\, ds$$
$$- \int_t^T \phi_1{}^*(s, t, \theta)\, dw(s) \qquad \text{for } 0 \leqslant t \leqslant T, \quad -b < \theta < 0 \quad (4.27)$$

where ϕ_0 and ϕ_1 are defined by (4.3) and (4.4). We set $p_1(T, \eta) = z(\eta)$ a.e. on $]-b, 0[$. Thus Eq. (4.26) can be written as

$$p(t) = \phi_0{}^*(T, t)\, p(T) - \int_t^T \phi_0{}^*(s, t)\, q(s)\, ds - \int_t^T \phi_0{}^*(s, t)\, dw(s)$$
$$+ \begin{Bmatrix} \int_{T-h}^{T} \phi_0{}^*(s, t)\, z(s - T)\, ds, & 0 < t < T - b \\ \int_t^T \phi_0{}^*(s, t)\, z(s - T)\, ds, & T - b < t < T \end{Bmatrix} \quad (4.28)$$

while

$$p_1(0, \theta) = \sum_{j=1}^{N} \begin{cases} A_j{}^*(\theta - \theta_j)\, p(\theta - \theta_j), & \theta_j < \theta < 0 \\ 0, & -b < \theta < \theta_j \end{cases}$$
$$+ \int_{-b}^{0} A_{01}^*(\theta - \eta, \eta)\, p(\theta - \eta)\, d\eta \quad (4.29)$$

and

$$\tilde{p}(0) = [p(0), p_1(0, \theta)], \qquad \tilde{p}(T) = [p(T), z(\theta)]. \quad (4.30)$$

We set

$$A_T{}^*(t)\, p(t) = A_{00}^* p(t) + \sum_{i=1}^{N} \begin{cases} A_i{}^*(t - \theta_i)\, p(t - \theta_i), & t - \theta_i \leqslant T \\ 0, & t - \theta_i > T \end{cases}$$
$$+ \begin{Bmatrix} \int_{-b}^{0} A_{01}(t - \theta, \theta)\, p(t - \theta)\, d\theta, & 0 \leqslant t \leqslant T - b \\ \int_{t-T}^{0} A_{01}(t - \theta, \theta)\, p(t - \theta)\, d\theta, & T - b \leqslant t \leqslant T \end{Bmatrix}. \quad (4.31)$$

Thus, Eq. (4.28) shows that the function $p(t)$ belongs to $BV(0, T; R^n)$ and satisfies the *hereditary adjoint* system

$$p'(t) + A_T^*(t)\, p(t) + \begin{cases} 0, & 0 < t < T - b \\ z(t - T), & T - b < t < T \end{cases} = q(t) + w' \qquad \text{on} \quad]0, T[\tag{4.32}$$

where p' and w' are taken in the sense of distributions.

Summarizing to this point, we have seen that according to definition given in Section 1, the pair (x, u) is said to be extremal for the given optimal control problem, iff there exist functions $p \in BV(0, T; R^n)$, $w \in BV(0, T; R^n)$, $q \in L^2(0, T; R^n)$ and $z \in L^2(-b, 0; R^n)$ satisfying Eq. (4.32) and

$$w' \in \mathcal{N}(x, K) \tag{4.33}$$

$$B^*(t)\, p(t) \in \partial_u L(x(t), u(t)); \qquad q(t) \in \partial_x L(x(t), u(t)), \quad \text{a.e.} \quad \text{on} \quad]0, T[\tag{4.34}$$

$$[\tilde{p}(0), -\tilde{p}(T)] \in \partial l(x_0, x_T) \tag{4.35}$$

where $\tilde{p}(0), \tilde{p}(T) \in M^2(-b, 0; R^n)$ are defined by (4.30).

Let $N(x, K)$ denote the *cone of normals* in R^n to K at the point x. This is the closed convex cone defined by

$$N(x, K) = \{y \in R^n; \langle y, x - u\rangle \geqslant 0 \text{ for all } u \in K\}.$$

Observe that the derivative p' in Eq. (4.32) is an R^n-valued measure on $[0, T]$. Thus we can write p' as

$$p' = \dot{p}(t)\, dt + \mu_s$$

where μ_s is a certain singular measure and $\dot{p}(t)$ is the ordinary derivative of $p(t)$ (which exists for almost every t). Similarly, the measure $\mu = w'$ can be expressed as

$$\mu = \mu_a(t)\, dt + \mu_s$$

where $\mu_a \in L^1(0, T)$ is the absolutely continuous part of μ. Then condition (4.33) can be expressed in the following equivalent form (see [16])

$$\mu_a(t) \in N(x(t), K) \quad \text{a.e.} \qquad t \in]0, T[$$

while

$$d\mu_s(t)/d\nu \in N(x(t), K) \qquad \nu \text{ a.e.} \tag{4.36}$$

where $d\mu_s/d\nu$ is the Radon–Nikodym derivative of μ_s and ν is any positive measure on $[0, T]$ with respect to which μ_s is absolutely continuous. A measure μ_s satisfying condition (4.36) is said to be *$N(x(t), K)$-valued.*

Then Eqs. (4.32) and (4.33) can be written under the following equivalent form

$$\dot{p}(t) + A_T^*(t)\,p(t) - \begin{cases} q(t), & 0 < t < T - b \\ q(t) - z(T - t), & T - b < t < T \end{cases} \in N(x(t), K)$$
$$\text{a.e.} \qquad t \in]0, T[. \qquad (4.37)$$

the singular part μ_s of $p' = dp$ *is* $N(x(t), K)$*-valued.* (4.38)

In particular, if there are no state constraints, then $N(x(t), K) = \{0\}$ and the dual extremal arc $p(t)$ is absolutely continuous.

Now we shall examine the circumstances under which Hypotheses (A), (B), (C) in Theorem 1, are satisfied in the particular case we considered.

Hypothesis (A) will be satisfied, assuming that

(A′) $\mathrm{Dom}_1\, H = R^n$, $0 \in \mathrm{int}\, \mathrm{Dom}_2\, H$,

where $H(x, p)$: $R^n \times R^m \to [-\infty, +\infty]$ is given by

$$H(x, p) = \sup\{\langle p, v\rangle - L(x, v);\ v \in R^m\},$$

and $\mathrm{Dom}\, H = \mathrm{Dom}_1\, H \times \mathrm{Dom}_2\, H$ is defined as in Section 1.

By a *feasible arc* in problem (4.9)–(4.11), we shall mean an $x \in W(-b, T; R^n)$ satisfying Eq. (4.1) condition (4.2) and

$$L(x, u) \in L^1(0, T),\ x(t) \in K \qquad \text{for } t \in [0, T]$$

where h and u a certain elements in $M^2(-b, 0; R^n)$ and $L^2(0, T; R^m)$, respectively. The pair $\{y^1, y^2\} \in M^2(-b, 0; R^n) \times M^2(-b, 0; R^n)$ is said to be attainable if there is at least one feasible arc x such that

$$x_0 = y^1, \qquad x_T = y^2.$$

As in general case, the set of all attainable pairs for given optimal control problem, will be denoted by C_L.

Then, Hypotheses (B) and (C) are satisfied under the following conditions:

(B′) *There is at least one feasible arc* x *such that*

$$l(x_0, x_T) < +\infty; \qquad x(t) \in \mathrm{int}\, K \qquad \textit{for every } t \in [0, T].$$

(C′) *There exists* $(x_0, x_T) \in C_L \cap D(l)$ *such that one of the following two conditions holds:*

$$x_T \in \mathrm{int}\{y \in M^2(-b, 0; R^n);\ (x_0, y) \in C_L\} \qquad (4.39)$$

$$x_T \in \mathrm{int}\{y \in M^2(-b, 0; R^n);\ (x_0, y) \in D(l)\}. \qquad (4.40)$$

In particular, condition (4.39) will be satisfied in the case in which

$$K = R^n, \qquad \mathscr{A}_h = \{x_T(\cdot, h, u); u \in L^2(0, T; R^n)\} = M^2(-b, 0; R^n)$$

and

$$L(x, u) \in L^1(0, T) \qquad \text{for every } \{x, u\} \in H^1(0, T; R^n) \times L^2(0, T; R^m).$$

Theorem 1 can therefore be applied to the present situation.

THEOREM 2. *Let Hypotheses* (A′), (B′) *and* (C′) *hold. Then the pair* $(x, u) \in W(-b, T; R^n) \times L^2(0, T; R^m)$ *is optimal in the problem* (4.9)–(4.11), *if and only if there are functions* $p \in BV(0, T; R^n)$, $q \in L^2(0, T; R^n)$ *and* $z \in L^2(T - h, T; R^n)$ *satisfying equations* (4.34, 4.35, 4.37) *and* (4.38).

Remark 1. The preceding argument shows that Theorem 1 is applicable to a more general class of optimization problems associated with hereditary differential system (4.1). Namely, defining

$$\tilde{L}(h, u) = \sum_{i=1}^{m} L_i(h(\tau_i), u) + \int_{-b}^{0} L_0(h(\theta), u)\, d\theta, \qquad \text{for } h \in M^2(-b, 0; R^n), \quad u \in R^m$$

where

$$0 = \tau_1 > \tau_2 > \cdots \tau_m = -a;$$

$$L_i\colon R^n \times R^m \to]-\infty, +\infty]; \qquad i = 0, 1,..., m,$$

we can express problem (4.12)–(4.14) as

$$\text{Minimize} \int_0^T \left(\sum_{i=1}^{m} L_i(x(t + \tau_i), u(t)) + \int_{-b}^{0} L_0(x(t + \theta), u(t))\, d\theta \right) dt + l(x_0, x_T) \tag{4.40}$$

in $x \in W(-b, T; R^n)$ and $u \in L^2(0, T; R^m)$ satisfying Eq. (4.10) and state constraint (4.11).

Under appropriate convexity and growth conditions on the functions L_i, Theorem 1 can be applied to extend the above optimality theorem to the present case.

Remark 2. In some particular cases the *interior* in Hypothesis (C′) can be replaced by the *interior* relative to a certain linear closed manifold in the space $M^2(-b, 0; R^n)$. This is the case of optimal control problems involving differential systems of the form (4.1) where the trajectories $x(t)$ must satisfy initial and terminal conditions such as $x_0 = \varphi$ and $x_T = \zeta$ where φ and ξ are given in $M^2(-b, 0; R^n)$ (see [1, 5]).

References

1. H. T. Banks and M. Q. Jacobs, An attainable sets approach to optimal control of functional differential equations with function space terminal conditions, *J. Differential Eqs.* **14** (1973), 127–149.
2. H. T. Banks and G. A. Kent, Control of functional differential equations of retarded type and neutral type to target sets in function space, *SIAM J. Control* **10** (1972), 567–594.
3. V. Barbu, Convex control problems of Bolza in Hilbert space, *SIAM J. Control* 13 (1975), 3.
4. V. Barbu, On convex control problems of Bolza in Hilbert space, *SIAM J. Control* **13** (1975), 4.
5. V. Barbu, Convex control problems for linear differential systems of retarded type, to appear.
6. H. Brezis, Opérateurs maximaux monotones et semigroupes de contraction dans les espaces de Hilbert, *in* "Math. Studies 5," North-Holland, 1973.
7. R. Datko, A linear control problem in an abstract Hilbert space, *J. Differential Eqs.* **9** (1971), 346–359.
8. R. Datko, Unconstrained control problems with quadratic cost, *SIAM Control J.* **11** (1973), 32–52.
9. M.-C. Delfour and S. K. Mitter, Controlability, observability and optimal feedback control of affine hereditary differential systems, *SIAM J. Control* **10** (1972), 298–328.
10. J. K. Hale, "Functional Differential Equations," Springer-Verlag, New York, 1971.
11. A. Halanay, On a boundary-value problem for linear systems with time-lag, *J. Differential Eqs.* **2** (1966), 47–56.
12. J. L. Lions, "Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles" Dunod-Gauthier-Villars, Paris, 1968.
13. R. T. Rockafellar, "Convex Analysis," Princeton Univ. Press, Princeton, N. J., 1970.
14. R. T. Rockafellar, Saddle-points and convex analysis, *in* "Differential Games and Related Topics" (Kuhn and Szego, Eds.), North-Holland, Amsterdam, London, 1971.
15. R. T. Rockafellar, Conjugate convex functions in optimal control and the calculus of variations, *J. Math. Anal. Appl.* **32** (1970), 174–222.
16. R. T. Rockafellar, State constraints in convex control problems of Bolza, *SIAM Control J.* **10** (1972), 691–715.